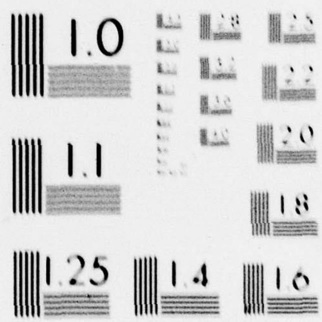
AD-A074 839 AEROSPACE CORP EL SEGUNDO CA ENGINEERING GROUP F/G 9/4
AN IMAGE TRANSFORM CODING ALGORITHM BASED ON A GENERALIZED CORR--ETC(U)
AUG 79 S M MELZER F04701-78-C-0079
UNCLASSIFIED TR-0079(4901-03)-6 SAMSO-TR-79-58 NL

| OF |

AD
A074839

END
DATE
FILMED
11-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

REPORT SAMSO-TR-79-58

(12)

LEVEL II

AD A074839

An Image Transform Coding Algorithm Based on a Generalized Correlation Model

S. M. MELZER
Engineering Group
The Aerospace Corporation
El Segundo, Calif. 90245

24 August 1979

Interim Report

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

Prepared for

SPACE AND MISSILE SYSTEMS ORGANIZATION
AIR FORCE SYSTEMS COMMAND
Los Angeles Air Force Station
P.O. Box 92960, Worldway Postal Center
Los Angeles, Calif. 90009

This interim report was submitted by The Aerospace Corporation, El Segundo, CA 90245, under Contract F04701-78-C-0079 with the Space and Missile Systems Organization (AFSC), Los Angeles Air Force Station, P.O. Box 92960, Worldway Postal Center, Los Angeles, CA 90009. It was reviewed and approved for The Aerospace Corporation by D. J. Griep, Engineering Group. Lieutenant J. C. Garcia, DYXT, was the Deputy for Technology project engineer.

This report has been reviewed by the Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication. Publication of this report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

James C. Garcia

James C. Garcia, Lt, USAF
Project Engineer

Joseph J. Cox Jr.

Joseph J. Cox, Lt Col
Chief, Advanced Technology Division

FOR THE COMMANDER

Burton H. Holaday

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

per inside cover + rept nr.

1. ~~REPORT NUMBER~~ 18 SAMSO TR-79-58	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	9 Technical rept.,
4. TITLE (and Subtitle) 6 AN IMAGE TRANSFORM CODING ALGORITHM BASED ON A GENERALIZED CORRELATION MODEL.	5. ~~TYPE OF REPORT & PERIOD COVERED~~ Interim
	6. ~~PERFORMING ORG. REPORT NUMBER~~ 14 TR-0079(4901-03)-6
7. AUTHOR(s) 10 Stuart M. Melzer	8. ~~CONTRACT OR GRANT NUMBER(s)~~ 15 F04701-78-C-0079
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Aerospace Corporation El Segundo, Calif. 90245	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Space and Missile Systems Organization/DYXT P.O. Box 92960, Worldway Postal Center Los Angeles, Calif. 90009	12. REPORT DATE 11 24 Aug[redacted] 1979
	13. NUMBER OF PAGES 12 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image
Transform
Coding
Statistics
Bits

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Image transform coding is a technique whereby a sampled image is divided into blocks. A two-dimensional discrete transform of each block is taken, and the resulting transform coefficients are coded. Coding of the transform coefficients requires their quantization and, consequently, a model for the transform coefficient variances that is based in turn on a correlation model for the image blocks. In the proposed correlation model each block of image data is formed by an arbitrary left and right matrix multiplication of a stationary white matrix. One consequence of this correlation model is that the

next page

19. KEY WORDS (Continued)

20. ABSTRACT (Continued)

transform coefficient variances are product separable in row and column indexes. The product separable model for the transform coefficient variances forms the basis of a transform coding algorithm. The algorithm is described and tested on real sampled images.

PREFACE

The author would like to thank Drs. R. V. Cox and A. G. Tescher for the use of their input-output subroutines and Mr. C. H. Garcia for his programming support.

This report was prepared as a paper for publication in the Proceedings of the Society of Photo-Optical Instrumentation Engineers, Vol. 149, Application of Digital Image Processing, August 1978.

CONTENTS

FIGURES

TABLES

AN IMAGE TRANSFORM CODING ALGORITHM BASED ON A GENERALIZED CORRELATION MODEL*

Introduction

Transform coding is an efficient technique for compressing coded image arrays from a high number to a low number of bits per pixel because the transform concentrates the image energy in the lower index transform coefficients. The developed algorithm discussed in this paper is within the framework of statistical, adaptive transform coding in that the transform coefficients are quantized with a quantizer and bit allocation based on the transform coefficient variances. The term adaptive transform coding is applicable because the transform coefficient variances for each block are estimated from the transform coefficients of that block. Recomputation of the transform coefficient variance estimates for each block allows these variance estimates to adapt to the statistical variations between blocks. This latter property is also found in the recursive estimation of the transform coefficient variances as proposed by Tescher and Cox.(1) Adaptive image coding techniques have been recently surveyed by Habibi.(2) The algorithm developed in this paper is new in the choice of model for the transform coefficient variances.

Statistical Model Theory

An image array is divided into $N \times N$ matrices called blocks. For the purpose of generating second-order statistics, the fundamental assumption made in this paper is that each block is formed by an outer product matrix multiplication on a zero mean stationary white matrix:

$$X = HWG^T \quad ; \quad \text{Trace}\,(H^TH) = \text{Trace}\,(G^TG) = N \tag{1}$$

In Eq. (1), X is the $N \times N$ matrix of image data for one block, H and G are $N \times N$ scaled matrices, and W is an $N \times N$ stationary white matrix.

$$E[w^2(k,i)] = \sigma^2 \quad ; \quad E[w(k,i)w(r,s)] = 0 \text{ for } k,i \neq r,s \tag{2}$$

In Eq. (2), $w(k,i)$ is the k, i element of W and $E[\cdot]$ is the statistical expectation operator.

The model proposed here is an extension of the model typically assumed for generating the second-order statistics of random vectors, i.e., that the random vector is formed by a matrix multiplication on a stationary white vector. The best known special case that fits the model [Eq. (1)] is that X is a block from a two-dimensional exponentially correlated image.

It follows from Eq. (2) that, for an arbitrary $N \times N$ matrix Q,

$$E[W^TQW] = E[WQW^T] = (\text{Trace}\;Q)\sigma^2 I \tag{3}$$

where I is the $N \times N$ identity matrix. It follows from Eqs. (1) and (3) that

$$\sigma^2 = \text{Trace}\,(E[XX^T/N^2]) \tag{4}$$

$$E[XX^T/N\sigma^2] = HH^T \quad ; \quad E[X^TX/N\sigma^2] = GG^T \tag{5}$$

*This work reflects research conducted under U.S. Air Force Space and Missile Systems Organization (SAMSO) Contract No. F04701-77-C-0078.

The matrix $E[XX^T/N\sigma^2]$ reflects the row correlations in X, and the matrix $E[X^TX/N\sigma^2]$ reflects the column correlations in X.

The Transform Statistics

The transform of the matrix X is the $N \times N$ matrix Z:

$$Z = U^TXV \quad ; \quad U^{-1} = U^T \quad , \quad V^{-1} = V^T \tag{6}$$

In Eq. (6), U and V are arbitrary unitary matrices and it follows that, given Z, then $X = UZV^T$. Subsequently, we investigate the second-order statistics of the transform coefficients $z(n, \ell)$. From (6)

$$z(n, \ell) = \underline{U}_n^T X \underline{V}_\ell \tag{7}$$

where $\underline{U}_n$ is the n^{th} column of U and $\underline{V}_\ell$ is the ℓ^{th} column of V. From Eqs. (1), (3), and (7), it follows that

$$E[z(n, \ell)z(r, s)] = \sigma^2\left(\underline{U}_n^T HH^T \underline{U}_r\right)\left(\underline{V}_s^T GG^T \underline{V}_\ell\right) \tag{8}$$

Substitution of Eq. (5) into Eq. (8) gives

$$E[z(n, \ell)z(r, s)] = \sigma^2\left(\underline{U}_n^T E[XX^T/N\sigma^2]\underline{U}_r\right)\left(\underline{V}_s^T E[X^TX/N\sigma^2]\underline{V}_\ell\right) \tag{9}$$

What is learned from Eqs. (4) and (9) is that the two matrices $E[XX^T]$ and $E[X^TX]$ constitute sufficient statistics for the determination of the second-order statistics of the $z(n, \ell)$ transform coefficients. Moreover, if $U^TE[XX^T]U$ is the unitary transformation that diagonalizes $E[XX^T]$ and if $V^TE[X^TX]V$ is the unitary transformation that diagonalizes $E[X^TX]$, then the transform coefficients are uncorrelated with $E[z(n, \ell)z(r, s)] = 0$ for $(n, \ell) \neq (r, s)$. In this case, $Z = U^TXV$ is the Karhunen-Loeve or Hotelling transformation for X.

From Eq. (9) it follows that the variance of the (n, ℓ) transform coefficient is

$$\sigma_{n,\ell}^2 \triangleq E[z^2(n, \ell)] = \sigma^2\lambda_n\beta_\ell \tag{10}$$

$$\lambda_n \triangleq \underline{U}_n^T E[XX^T/N\sigma^2]\underline{U}_n \quad ; \quad \beta_\ell \triangleq \underline{V}_\ell^T E[X^TX/N\sigma^2]\underline{V}_\ell \tag{11}$$

What is learned from Eq. (10) is that the transform coefficient variances are product separable in row and column indexes. Since XX^T and X^TX are positive semidefinite matrices and since U and V are unitary, it follows that

$$\lambda_n \geq 0 \quad , \quad \sum_{n=1}^{N}\lambda_n = N \quad ; \quad \beta_\ell \geq 0 \quad , \quad \sum_{\ell=1}^{N}\beta_\ell = N \tag{12}$$

From Eqs. (10) and (12), σ^2, λ_n, and β_ℓ can be expressed directly in terms of the $E[z^2(n, \ell)]$:

$$\sigma^2 = \sum_{\ell=1}^{N}\sum_{n=1}^{N} E[z^2(n, \ell)]/N^2 \tag{13}$$

$$\lambda_n = \sum_{\ell=1}^{N} E[z^2(n, \ell)]/N\sigma^2 \quad ; \quad \beta_\ell = \sum_{n=1}^{N} E[z^2(n, \ell)]/N\sigma^2 \tag{14}$$

The results of applying the rate distortion theory to random vectors[3] carry through analogously to random matrices modeled by Eq. (1). If the elements of W are Gaussian, then the least number of bits per element B required to block code X with mean square distortion D per element is

$$B = 0.5\log_2(\sigma^2/D) - (0.5/N)\left[\log_2\left(\mathrm{Det}R_1^{-1}\right) + \log_2\left(\mathrm{Det}R_2^{-1}\right)\right] \tag{15}$$

$$R_1 \triangleq E[XX^T/N\sigma^2] \quad \text{and} \quad R_2 \triangleq E[X^TX/N\sigma^2] \tag{16}$$

In Eq. (15), D/σ^2 is assumed to be less than the product of the smallest eigenvalue of R_1 times the smallest eigenvalue of R_2 with R_1 and R_2 restricted to be positive definite. The proof is discussed below with $b(k, i)$, $\hat{b}(k, i)$, $z(n, \ell)$ and $\hat{z}(n, \ell)$ denoting the elements of X, $\hat{X}$, Z and $\hat{Z}$ respectively and $\hat{X} = U\hat{Z}V^T$.

Rate Distortion Results

The block mean square distortion is defined as $D_T = \Sigma\Sigma E[(b(k,i) - \hat{b}(k,i))^2]$. Since U and V are unitary, it follows that $D_T = \Sigma\Sigma E[(z(n,\ell) - \hat{z}(n,\ell))^2]$. Each transform coefficient $z(n,\ell)$ is coded with $B_{n,\ell}$ bits and decoded as $\hat{z}(n,\ell)$ with mean square distortion $D_{n,\ell} = E[(z(n,\ell) - \hat{z}(n,\ell))^2]$. Thus, $D_T = \Sigma\Sigma D_{n,\ell}$. We next assume that the elements of W in the model given in (1) are Gaussian, and it follows that the $z(n,\ell)$ are Gaussian. The minimum bit allocation $B_{n,\ell}$ to achieve the mean square distortion $D_{n,\ell} \le \sigma^2_{n,\ell}$ for Gaussian $z(n,\ell)$ is $B_{n,\ell} = 0.5\log_2(\sigma^2_{n,\ell}/D_{n,\ell})$. The total bit allocation for the block is $B_T = \Sigma\Sigma B_{n,\ell}$ and it follows that

$$B_T = 0.5\Sigma\Sigma\log_2\sigma^2_{n,\ell} - 0.5\Sigma\Sigma\log_2 D_{n,\ell} \tag{17}$$

To minimize B_T for a given D_T, we choose $D_{n,\ell}$ to maximize $\Sigma\Sigma\log_2 D_{n,\ell}$. The solution is $D_{n,\ell} = D_T/N^2$ provided $D_T/N^2 \le \sigma^2_{n,\ell}$ for all indexes. Considering the case where this last inequality is satisfied, then

$$B_T = 0.5\Sigma\Sigma\log_2\sigma^2_{n,\ell} - 0.5N^2\log_2\left(D_T/N^2\right) \tag{18}$$

Denoting $D = D_T/N^2$ as the normalized distortion, $B = B_T/N^2$ as the normalized bit allocation, and substituting $\sigma^2_{n,\ell} = \sigma^2\lambda_n\beta_\ell$, it follows that

$$B = 0.5\log_2(\sigma^2/D) + (0.5/N)\left[\sum_{n=1}^{N}\log_2\lambda_n + \sum_{\ell=1}^{N}\log_2\beta_\ell\right] \tag{19}$$

From (11) and (16) the λ_n are the main diagonal terms of $U^T R_1 U$ and the β_ℓ are the main diagonal terms of $V^T R_2 V$. The term B is minimized when $\Pi\lambda_n$ and $\Pi\beta_\ell$, which are the respective products of these main diagonals, are minimized. This occurs when $Z = U^T X V$ is the Karhunen-Loeve Transform for X. Then the λ_n are the eigenvalues of R_1 and the β_ℓ are the eigenvalues of R_2. It follows that $\Sigma\log_2\lambda_n = \log_2(\Pi\lambda_n) = \log_2(\mathrm{Det}R_1)$ and similarly for R_2. Substitution into (19) gives

$$B = 0.5\log_2(\sigma^2/D) + (0.5/N)[\log_2(\mathrm{Det}R_1) + \log_2(\mathrm{Det}R_2)] \tag{20}$$

Equation (16) follows since $\mathrm{Det}R_1^{-1} = (\mathrm{Det}R_1)^{-1}$ and similarly for R_2. Also, the restriction on D is $D/\sigma^2 \le \lambda_n\beta_\ell$ for all n, ℓ.

Description of the Algorithm

The basic structure of statistical transform coding algorithms is given in Fig. 1. The original image is a large dimensional array whose elements are each coded using B_o bits, and the images in our study have elements that take on integer values from 0 to $2^{B_o} - 1$. The first step is to divide the image into $N \times N$ blocks with the choice in this study of $N = 16$. The 16×16 block size is large enough to take advantage of the compression of energy into the lower index transform coefficients and small enough to permit consideration of statistical variations over the image. Next, the two-dimensional discrete cosine transform operation $U = V$, $Z = U^T X U$ is performed. The elements $u(n,\ell)$ of the $N \times N$ unitary matrix U are

$$u(n,1) = 1/\sqrt{N} \quad ; \quad n = 1, 2 \cdots N \tag{21}$$

$$u(n,\ell) = \sqrt{\frac{2}{N}}\cos\left[\frac{(\ell-1)\pi(2n-1)}{2N}\right] \quad ; \quad \begin{matrix} n = 1, 2 \cdots N \\ \ell = 2, 3 \cdots N \end{matrix} \tag{22}$$

The cosine transform is a popular choice because of the closeness to the Karhunen-Loeve Transform for exponentially correlated images.[4] In this study no fast Fourier Transform-type algorithm is used. The cosine transform $Z = U^T X U$ and the inverse $\hat{X} = U\hat{Z}U^T$ are computed as written – by matrix multiplication.

Unlike the assumption made in the theoretical discussion, the real image blocks X do not have zero mean. For the two-dimensional cosine transform, only the $z(1,1)$ transform coefficient is biased. Using (21)

$$z(1,1) = \sum_{i=1}^{N}\sum_{k=1}^{N} b(k,i)/N \tag{23}$$

where $b(k,i)$ is an element of X. Since each $b(k,i)$ is an integer on 0 to $2^{B_o} - 1$, then $Nz(1,1)$ is an integer on 0 to $N^2(2^{B_o} - 1)$ that requires less than $2\log_2 N + B_o$ bits to code without error. In this study, $z(1,1)$ is decoded exactly (i.e., $\hat{z}(1,1) = z(1,1)$) with the required number of bits included in the bit count.

In the next step, the σ^2, λ_n, and β_ℓ transform statistics are estimated for each block using that block's transform coefficients to replace $E[z^2(n,\ell)]$ in (13) and (14):

$$(\sigma')^2 = \sum_{\ell=1}^{N} \sum_{n=1}^{N} z^2(n, \ell)/N^2 \tag{24}$$

$$\lambda'_n = \sum_{\ell=1}^{N} z^2(n, \ell)/N(\sigma')^2 \quad ; \quad \beta'_\ell = \sum_{n=1}^{N} z^2(n, \ell)/N(\sigma')^2 \tag{25}$$

In Eqs. (24) and (25), $(\sigma')^2$, λ'_n, and β'_ℓ denote the estimates of σ^2, λ_n, and β_ℓ. Also, the actual $z^2(1, 1)$ is replaced by zero.

The estimates of the block transform statistics must be coded since they are required for obtaining the decoded transform coefficients. Let $\hat{\sigma}^2$, $\hat{\lambda}_n$, and $\hat{\beta}_\ell$ denote the quantized $(\sigma')^2$, λ'_n, and β'_ℓ respectively. Then the transform coefficient quantized standard deviation is

$$\hat{\sigma}_{n,\ell} = \left[\hat{\sigma}^2 \hat{\lambda}_n \hat{\beta}_\ell\right]^{1/2} \tag{26}$$

Each $z(n, \ell)$ with the exception of $z(1, 1)$ is normalized by $\hat{\sigma}_{n,\ell}$ to obtain the normalized transform coefficient

$$\chi(n, \ell) = z(n, \ell)/\hat{\sigma}_{n,\ell} \tag{27}$$

The normalized transform coefficient is then quantized using a $\hat{B}_{n,\ell}$ bit unity variance Max-Lloyd[5] Gaussian quantizer. The choice of $\hat{B}_{n,\ell}$ is based upon $\hat{\sigma}_{n,\ell}$ as discussed later.

Decoding the bit word for the quantized normalized transform coefficient yields $\hat{\chi}(n, \ell)$, which is multiplied by $\hat{\sigma}_{n,\ell}$ to obtain the decoded transform coefficient

$$\hat{z}(n, \ell) = \hat{\sigma}_{n,\ell} \hat{\chi}(n, \ell) \tag{28}$$

The same $\hat{\sigma}_{n,\ell}$ must be used in (27) and (28), which is why $(\sigma')^2$, λ'_n, and β'_ℓ must be quantized and included in the bit stream. Finally, the matrix of decoded transform coefficients $\hat{Z} = [\hat{z}(n, \ell)]$ is inverted to obtain the decoded image block $\hat{X} = [\hat{b}(k, i)]$:

$$\hat{X} = U\hat{Z}U^T \tag{29}$$

Coding of the Transform Statistics

Since $\hat{\sigma}_{n,\ell}$ is used to normalize $z(n, \ell)$, accurate quantizing of the transform statistics means that $(\hat{\sigma}/\sigma')$, $(\hat{\lambda}_n/\lambda'_n)^{1/2}$, and $(\hat{\beta}_\ell/\beta'_\ell)^{1/2}$ are near 1. It follows that the transform statistics should be logarithmically quantized. For $(\hat{\sigma})^2$, the output choices used are 1, C, $C^2 \cdots C^{K-1}$ where K is the number of choices. The number of bits required to code $(\sigma')^2$ is $\log_2 K$ and is counted as overhead. For $(\sigma')^2$ between the smallest and largest output choices,

$$(\hat{\sigma})^2 = C^{\mathrm{INT}[\log_2(\sigma')^2/\log_2 C]} \tag{30}$$

where INT[ ] is used here as the closest integer to the argument.

The primary contribution to overhead is the coding of the 2N statistical coefficients, $\lambda'_1 \cdots \lambda'_N$ and $\beta'_1 \cdots \beta'_N$. If we allocated B bits to each coefficient, then the overhead is $2NB/N^2 = 2B/N$ bits per pixel. The coding of λ'_n for N=16 is as follows. From (24) and (25) $\lambda'_n \geq 0$ and $\Sigma\lambda'_n = 16$, and it follows that $0 \leq \lambda'_n \leq 16$. The eight output level choices of 1/16, 1/8, 1/4, 1/2, 1, 2, 4, and 8 are chosen for $\hat{\lambda}_n$ and the logarithmic quantization rule is

$$\hat{\lambda}_n = 2^{\mathrm{INT}[\log_2 \lambda'_n]} \qquad 1/16 < \lambda_n < 8$$

$$\hat{\lambda}_n = 1/16 \text{ for } \lambda'_n \leq 1/16 \text{ and } \lambda'_n = 8 \text{ for } \lambda'_n \geq 8 \tag{31}$$

The same approach is used to code the β'_ℓ.

If the same number of bits is allocated to each of the 8 choices, then each coefficient is allocated 3 bits and the resulting overhead for coding all 32 statistical coefficients is (2)(3)/16 = 0.375 bits per pixel. However, since the transform tends to concentrate block energy in the lower index transform coefficients, the $\hat{\lambda}_n$ obtained from (31) will tend to decrease with increasing index n. Thus, for each index n of $\hat{\lambda}_n$ there are some choices more likely to be selected than others. For example, $\hat{\lambda}_1$ is nearly always 4 or 8. The implication is that we can improve on the overhead of 0.375 bits per pixel for the 32 statistical coefficients by using a variable word length code.

A Fano code was tried that uses code words of 1 to 7 bits for the eight choices. The code words in the order of decreasing probability are 0, 1-0, 1-1-0, 1-1-1-0, 1-1-1-1-0, 1-1-1-1-1-0, 1-1-1-1-1-1-0, and 1-1-1-1-1-1-1. The most probable $\hat{\lambda}_n$ is coded 0, the second most probable is coded 1-0, and so on. Because the sequence $\hat{\lambda}_1 \cdots \hat{\lambda}_N$ tends to be correlated, the order of probability for $\hat{\lambda}_n$ was set from $\hat{\lambda}_{n-1}$ for n=2, 3 . . . 16 with the order fixed for $\hat{\lambda}_1$. The rule used for setting the order of probability is given in Table 1. The lowest overhead obtainable with this scheme is 0.125 bits per pixel for all 32 coefficients. There are other methods for quantizing the λ'_n that might be tried, including differential pulse code modulation on the sequence $\log_2\lambda'_1 \cdots \log_2\lambda'_{16}$, one-dimensional transform coding of this sequence, and curve fitting to a finite parameter model of this sequence.

Bit Allocation Rules for the Transform Coefficients

The number of bits $\hat{B}_{n,\ell}$ allocated to code $\chi(n,\ell)$ is determined from the bit allocation rule. A noninteger possible negative $B_{n,\ell}$ is computed and then set to the closest positive integer or zero:

$$\hat{B}_{n,\ell} = \mathrm{INT}[B_{n,\ell}] \text{ for } B_{n,\ell} > 0 \quad \text{and} \quad \hat{B}_{n,\ell} = 0 \text{ for } B_{n,\ell} \le 0 \tag{32}$$

The rules apply to all $\chi(n,\ell)$ with the exception of $n, \ell = 1, 1$. As discussed earlier, z(1, 1) is error-free coded using $B_{1,1} = 2\log_2 N + B_o = 8 + B_o$ bits. For the $\chi(n,\ell)$, if $\hat{B}_{n,\ell} = 0$, then $\hat{\chi}(n,\ell) = 0$ is decoded.

The first allocation rule is called here "local bit control" (LBC) and is used to allocate approximately the same number of bits to each block. The rule is

$$B_{n,\ell} = B_T/N^2 - (0.5/N^2)\left[\sum_{\ell=1}^{N}\sum_{n=1}^{N}\log_2\hat{\sigma}^2_{n,\ell}\right] + 0.5\log_2\hat{\sigma}^2_{n,\ell} \tag{33}$$

Note that $\Sigma\Sigma B_{n,\ell} = B_T$ but, since $\hat{B}_{n,\ell} = 0$ for $B_{n,\ell} < 0$, then $\Sigma\Sigma\hat{B}_{n,\ell}$ will be somewhat greater than B_T. This rule tends to keep the allocations for the blocks about the same, but results in block distortions that differ from block to block reflecting their statistical variations.

The second allocation rule is called here "global distortion control" (GDC) and is used to maintain approximately the same block distortions for each block. This rule is

$$B_{n,\ell} = 0.5\log_2(\hat{\sigma}^2_{n,\ell}/D) \tag{34}$$

Note that $\Sigma\Sigma B_{n,\ell} = 0.5\Sigma\Sigma\log_2\hat{\sigma}^2_{n,\ell} - 0.5N^2\log_2 D$. The bit allocation will vary between blocks, reflecting the differences in block quantized transform statistics.

Performance Evaluation

Performance Equations

The performance of the algorithm is evaluated using a normalized mean square error (MSE) criterion. The MSE for the entire image is defined by

$$\mathrm{MSE} \triangleq \frac{\mathrm{DST}}{\mathrm{ENGY}} \tag{35}$$

where DST is the distortion per pixel and ENGY is the energy per pixel. The distortion per pixel is defined by

$$\mathrm{DST} \triangleq \frac{1}{N^2M}\sum_{r=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{N}(b_r(k,i) - \hat{b}_r(k,i))^2 \tag{36}$$

where r is the block index out of M total blocks, $b_r(k,i)$ is the input image, and $\hat{b}_r(k,i)$ is the output image. Since the transform is unitary, DST can be expressed in terms of the transform coefficients by

$$\mathrm{DST} = \frac{1}{N^2M}\sum_{r=1}^{M}\sum_{\ell=1}^{N}\sum_{n=1}^{N}(z_r(n,\ell) - \hat{z}_r(n,\ell))^2 \tag{37}$$

The energy per pixel is defined by

$$\mathrm{ENGY} \triangleq \frac{1}{N^2M}\sum_{r=1}^{M}\sum_{i=1}^{N}\sum_{k=1}^{N}(b_r(k,i) - m)^2 \tag{38}$$

where m is the image global sample mean.

$$m \triangleq \frac{1}{N^2 M} \sum_{r=1}^{M} \sum_{i=1}^{N} \sum_{k=1}^{N} b_r(k, i) \qquad (39)$$

Note from (35) and (38) that the MSE is computed here with the global sample mean extracted.

The bits per pixel that is used to achieve the MSE is $B_A + B_S$ where B_A is the bits per pixel used to code the transform coefficients and B_S is the bits per pixel used to code the transform statistics. The relation for B_A is

$$B_A = \frac{1}{N^2 M} \sum_{r=1}^{M} \sum_{\ell=1}^{N} \sum_{n=1}^{N} \hat{B}_{n,\ell}(r) \qquad (40)$$

In (40), $\hat{B}_{n,\ell}(r)$ is the bit allocation for the n, ℓ transform coefficient of block r. As discussed earlier, $\hat{B}_{1,1}(r) = 8 + B_o$ bits and all other $\hat{B}_{n,\ell}(r)$ follow from (32) and (33) or (34). The relation for B_S is

$$B_S = \frac{1}{N^2 M} \sum_{r=1}^{M} \left[\text{block r bits used to code } (\sigma')^2, \lambda'_1 \cdots \lambda'_N \text{ and } \beta'_1 \cdots \beta'_N\right] \qquad (41)$$

Test Results

The algorithm is tested on two satellite weather images and a site image. Image specifications are given in Table 2, and performance results for MSE compared with the bits per pixel $B_A + B_S$ are presented in Table 3. Selected original, reconstructed, and difference images are shown in Figs. 2, 3, and 4.

The contribution to B_S as the result of coding the $(\sigma')^2$ is $\log_2 K/256$ bits per pixel, which is 0.019 for the weather images (i.e., K = 32, C = 1.25) and 0.023 for the site image (i.e., K = 64, C = 1.22). The contribution to B_S is 0.375 bits per pixel as the result of coding the λ's and β's without a variable word length code. Simulated reductions were obtained using the Fano code of Table 1. For weather image 1, the contribution to B_S was reduced to 0.285; for weather image 2, to 0.307; and for the site image, to 0.277. These results are reflected in Table 3.

For a given bits per pixel, global distortion control provides a lower mean square error than local bit control as may be seen by comparing the first and second cases for weather image 2 given in Table 3. A characteristic of local bit control is that the difference image will have "blocking" reflecting the different distortions on blocks with different statistics. This effect can be seen in the difference image shown in Fig. 2 for weather image 1.

In Figs. 2, 3, and 4 the reconstructed images are, to the eye, replications of the originals. Thus, to accentuate differences between the original and reconstructed images, an amplified difference image is shown representing $30|b(k, i) - \hat{b}(k, i)|$ with white corresponding to higher difference than black. Note that the difference images in Figs. 3 and 4 have the pepper appearance that indicates the fairly uniform distortion over the image, which is characteristic of global distortion control.

Conclusions

A transform coding algorithm has been developed based on a generalized correlation model of the image blocks. The algorithm requires coding of the transform coefficients plus the overhead, which is the bits per pixel required to code the transform statistics. The developed algorithm is characterized by a low mean square error between original and reconstructed images for the bits per pixel allocated to code the transform coefficients without excessive overhead. Among various possible schemes for coding the transform statistics, an eight-level quantizer followed by a variable word length code was utilized. For all tests the reconstructed images are very good, supporting the reasonableness of the product separable model for the transform coefficient variances. The model for the transform coefficient variances is sufficiently generalized so that high frequency image structure can be reflected. Computation of the quantized transform coefficient variances prior to coding the transform coefficients permits control of the bit rate since the number of bits allocated to each block is known prior to coding.

References

1. Tescher, A. G. and Cox, R. V., "An Adaptive Transform Coding Algorithm," Proc. of the XII IEEE International Conference on Communications, June 1976.

2. Habibi, Ali, "Survey of Adaptive Image Coding Techniques," IEEE Trans. on Communications, Vol. Com-25, No. 11, pp. 1275-1283, November 1977.

3. Davidson, L. D., "Rate Distortion Theory and Application," IEEE Proc. - Special Issue on Digital Picture Processing, Vol. 60, No. 7, pp. 800-808, July 1972.
4. Ahmed, N., Natarajan, T., and Rao, K. R., "Discrete Cosine Transform," IEEE Trans. on Computers, Vol. C-23, pp. 90-93, January 1974.
5. Max, J., "Quantizing for Minimum Distortion," IRE Trans. on Information Theory, Vol. IT-6, pp. 7-12, March 1960.

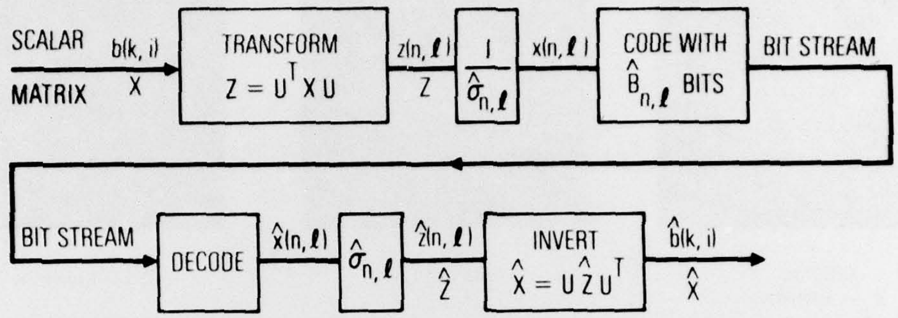


Fig. 1. Basic structure of statistical transform coding.

ORIGINAL
$B_0 = 6$ Bits/Pixel

RECONSTRUCTED

ABS (Difference) × 30

MSE × 100 = 0.23%

$B_A = 1.575$ Bits/Pixel - Transform Coefficients

$B_S = 0.304$ Bits/Pixel - Transform Statistics

Fig. 2. Local bit control for weather image 1.

ORIGINAL
$B_0 = 6$ Bits/Pixel

RECONSTRUCTED

ABS (Difference) × 30

MSE × 100 = 1.65%

$B_A = 0.256$ Bits/Pixel Transform Coefficients

$B_S = 0.326$ Bits/Pixel Transform Statistics

Fig. 3. Global distortion control for weather image 2.

ORIGINAL
$B_0 = 10$ Bits/Pixel

RECONSTRUCTED

ABS (Difference) × 30

MSE × 100 = 0.34%

$B_A = 2.102$ Bits/Pixel Transform Coefficients

$B_S = 0.300$ Bits/Pixel Transform Statistics

Fig. 4. Global distortion control for site image.

Table 1. Bit Allocation Rule for Coding of Statistical Coefficients Using Fano Code

	1 bit	2 bit	3 bit	4 bit	5 bit	6 bit	7 bit	7 bit
$\hat{\lambda}_1$	8	4	2	1	1/2	1/4	1/8	1/16
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 8$	4	2	8	1	1/2	1/4	1/8	1/16
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 4$	2	1	4	1/2	8	1/4	1/8	1/16
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 2$	1	1/2	2	1/4	4	1/8	8	1/16
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 1$	1/2	1/4	1	1/8	2	1/16	4	8
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 1/2$	1/4	1/8	1/2	1/16	1	2	4	8
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 1/4$	1/8	1/16	1/4	1/2	1	2	4	8
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 1/8$	1/16	1/8	1/4	1/2	1	2	4	8
$\hat{\lambda}_n$ given $\hat{\lambda}_{n-1} = 1/16$	1/16	1/8	1/4	1/2	1	2	4	8

Table 2. Image Specifications

	B_o Bits/Pixel	Image Dimensions	No. of Blocks(M)	Global Sample Mean(m)	Energy per Pixel (ENGY)
Weather Image 1	6	480 × 704	1320	30.76	191.26
Weather Image 2	6	480 × 704	1320	25.88	23.70
Site Image	10	512 × 512	1024	306.00	7269.

Table 3. Performance Results

	B_A	B_S	$B_A + B_S$	MSE × 100 (%)	Bit Allocation Rule
Weather Image 1 $B_o = 6$	0.187	0.304	0.491	1.47	GDC
	0.827	0.304	1.131	0.50	GDC
	1.575	0.304	1.879	0.23	LBC
Weather Image 2 $B_o = 6$	0.256	0.326	0.582	1.65	GDC
	0.306	0.326	0.632	1.82	LBC
	0.631	0.326	0.957	0.91	GDC
Site Image $B_o = 10$	0.215	0.300	0.515	4.53	GDC
	0.562	0.300	0.862	2.44	GDC
	2.102	0.300	2.402	0.34	GDC

B_o = Original Bits/Pixel

B_A = Bits/Pixel Transform Coefficients

B_S = Bits/Pixel Overhead for Transform Statistics with Fano Code

LBC = Local Bit Control

GDC = Global Distortion Control